DOCUMENT RESUME

ED 370 617 JC 940 311

AUTHOR Isonio, Steven
TITLE Relationship between APS Writing Test Scores and Instructor Preparedness Ratings: Further Evidence for Validity.
INSTITUTION Golden West Coll., Huntington Beach, Calif.
PUB DATE Jun 94
NOTE 15p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Faculty; Community Colleges; *Evaluation Problems; Grades (Scholastic); Predictive Validity; *Predictor Variables; *Screening Tests; *Student Placement; Test Validity; Two Year Colleges; *Writing Skills; *Writing Tests

ABSTRACT

In fall 1993, Golden West College (GWC), in California, participated in a large-scale study coordinated by Computerized Assessment and Placement Programs (CAPP) to gather evidence supportive of the College Board Assessment and Placement Services (APS) tests. The study focused on gathering instructor ratings of student preparedness levels in a range of courses to serve as the criterion for an evaluation of the validity of the APS Writing test, GWC's primary instrument for English placement. Using end-of-course grades as the criterion in traditional predictive validity studies assumes that skills measured by the test are necessary for student success. Yet, a wide array of factors combine to determine course grades and educational success. A more appropriate gauge for assessment is whether the tests provide meaningful information about the students' skill levels. Instructors in English writing courses provided ratings of the level of preparedness of each of their students. A total of 312 ratings were then matched with CAPP assessment data. The primary analysis involved determining the correlation between APS writing test scores and instructor ratings. The findings supported the conclusion that the APS Writing test is an appropriate and effect placement tool and provided indirect evidence for the multiple-measure placement model currently used at GWC. The data also supported the argument that instructor ratings can be effective criteria in placement validation studies. Descriptions of the targeted English writing classes, directions for instructor ratings of student preparedness, and correlation data are appended. (KP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 370 617

Relationship Between APS Writing Test Scores and Instructor Preparedness Ratings: Further Evidence for Validity

GOLDEN WEST COLLEGE
Huntington Beach, CA

Steven Isonio, Ph.D.
Advisor, Matriculation Research & Assessment

June 1994

JC 940 311

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. Isonio

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

Relationship Between APS Writing Test Scores and Instructor Preparedness Ratings: Further Evidence for Validity

Background

The College Board APS Writing test (Form A), a 40-item multiple-choice instrument, was adopted as the primary placement instrument for English placement at Golden West College (GWC) in Spring of 1991. As a part of the original implementation study, data were collected to establish initial cutoff scores and to evaluate the validity of that test for local placement purposes (Isonio, 1991). In the Spring of 1993, a multiple measures model which incorporates APS Writing test scores and information from key background questions (high school grade-point-average and number of years of English instruction in high school) was developed, adopted, and implemented. Although colleges are required to use multiple measures for placement (California Community Colleges, 1992), minimal standards for validity still apply to individual measures. This report presents further evidence for the validity of the APS Writing test at GWC.

In the Fall 1993 semester, Golden West College was given an opportunity to participate in a large scale study coordinated by Computerized Assessment and Placement Programs (CAPP) designed to gather evidence supportive of the College Board Assessment and Placement Services tests (CAPP Associates, 1993). The focus

of the study was on gathering ratings by instructors of the level of student preparedness for a range of courses. These ratings served as the criterion for an evaluation of the validity of the APS Writing test, the primary English placement instrument at GWC.

Appropriateness of *Placement* Validity.

There is growing recognition of the many problems associated with the use of end-of-course grades as the criterion in traditional predictive validity studies. The use of grades as the criterion in such studies assumes that the skills measured by the placement tests are both sufficient and necessary for success in the class. Yet, a wide array of factors, many of which are non-academic, combine to determine course grades specifically, and educational success generally (e.g., Woehr & Cavell, 1993; Cabrera, Nora & Casteneda, 1993). Thus, the notion that placement tests, instruments that were never intended to measure these direct and interacting factors, can reasonably be expected to serve as effective predictors of grades is being seriously questioned. A more appropriate gauge for the assessment of placement tests is whether they provide meaningful information about, and support correct decisions regarding, students' skill levels. This is the issue of student preparedness.

Traditional predictive validity coefficients are often difficult to interpret. They likely underestimate the true strength of the relationship between predictor and criterion since these indices reflect an assortment of factors that are independent of

the true underlying relationship. Constrained distributions, as would occur if an analysis is based upon data for a subgroup of students who had placed into the course by means of the test, result in attenuation of the correlation coefficient. Finally, and perhaps more important, since validity is limited by reliability, any degree of unreliability of measures of either predictor or criterion--test scores or course grades--will also attenuate the obtained correlation coefficient. There is considerable evidence for instructor grading variability and its impact on predictive validity evaluations (Rasor & Barr, 1993).

Beyond the array of methodological and analytic pitfalls associated with it, there are a number of conceptual reasons for not limiting the examination of placement validity to the traditional predictive model. In addition to the statistical artifacts related to predictive validity outlined above, the proper use of placement tests is placement--interpreting of test scores as indicators of *a current level of preparedness*. As such, the proper gauge of the effectiveness of placement instruments is their ability to place students in initial math, English, reading, and ESL courses for which the students are best suited. While doing so in an effective manner should indeed lead to increased student success, the key issue related to validity is not prediction. Rather, it reflects the tests' abilities to provide evaluations of existing skill levels. These evaluations, represented by test scores or composite indices, when properly articulated with course structures result in a match between skill levels of students and requisite entry skill levels of courses.

Kane's (1992) description of an argument-based approach to validity, with many illustrations from placement testing, is relevant here. He notes that there are usually many possible interpretations which can be reasonably assigned to a given set of test scores. As an example, a test consisting of series of questions following reading passages could be interpreted variously as a measure of the student's ability to read passages of that type and answer related questions, as a measure of general reading comprehension, as an indicator of overall verbal ability, or even much more broadly as a measure of intelligence. Not only is validity not an inherent characteristic of a test, but even the *interpretation* of a given test can reasonably vary. The implication of this for the consideration of placement validity is that there is a need to clearly specify the interpretation that is made of placement test scores. Interpreting them as predictors of course grades, as noted above, seems inappropriate. Other factors such as student motivation, personal problems, teacher styles, and grading practices are often greater determinants of student success as measured by final grades. The best interpretation of placement test scores is one that is closer to diagnosis than to prediction. As such, criterion measures such as ratings of readiness taken after a short period of instruction are typically better than end of course grades. The appropriateness of such "concurrent validation-like" evidence as Anastasi (1982) has argued "can be justified on its own right". She continues, "the logical distinction between predictive and concurrent validation is based, not on time, but on the objectives of testing. Concurrent validation is relevant to tests employed for diagnosis of existing states, rather than

prediction of future outcomes" (Anastasi, 1982, p. 137). Within the context of assessment and placement, "diagnosis of existing states" is synonymous with "determination of readiness".

Method

Instructors of English writing courses at GWC (English 9, English 10, and English 100--see Appendix A for descriptions of these courses) provided ratings of the level of preparedness of each of their students for the course in which they were enrolled. These ratings were matched with assessment data compiled by the CAPP software program in the GWC Assessment Center. A total of 312 such data matches were made. Care was taken to include only those students who had taken the APS Writing test during the testing period which immediately preceded the Fall 1993 semester. Cases where students had older test scores were considered inappropriate for inclusion in an analysis of the validity of the test. Skill levels may have changed for any of a variety of reasons including the possibility that the student may have had some instruction that would increase skill levels, or that without practice, skill levels might decrease over time.

The primary analysis involved computing the correlation between APS Writing test scores and the instructor ratings. These correlation coefficients can be interpreted as validity coefficients since they represent the relationship between the test scores and an appropriate criterion. More specifically, they can be properly

characterized as placement validity coefficients to distinguish them from predictive validity coefficients.

As a part of the large scale study for which this author was a principal investigator, similar data were collected at a number of community colleges in the state. At GWC, the data were collected near the midpoint of the Fall 1993 semester. Specifically, instructors were asked to weigh students' demonstrated ability to understand material covered in the course, based on direct observation and objective performance evaluations. Motivation and effort were not to influence the ratings. The focus was on evaluating student *preparedness* -- the goal of assessment/placement. A 5-point Likert scale, ranging from "unprepared for the course" to "exceptionally well prepared for the course" was used for the ratings (See Appendix B).

Results and Discussion

The primary analysis involved determining the correlation between the instructor ratings of student preparedness and student APS writing test scores. Table 1 (see Appendix C) presents both the obtained and disattenuated correlation coefficients. Since the coefficients were calculated separately for students in specific classes (English 9, English 10, and English 100 as the subgroups), the values are likely attenuated by the restricted range of scores associated with a particular course level. As such, the correction for restriction of range is appropriate and was applied

(Matriculation Local Research Options Committee, 1991). The corrected values, ranging from r=.30 for English 100 to a high of r=.53 for English 10, constitute strong evidence for the validity of the APS writing test as a placement instrument at Golden West College. These findings effectively corroborate the validation evidence gathered when the APS Writing test was piloted for implementation in early 1991, and support the conclusion that the APS Writing test is an appropriate and effective tool for placement of new students into initial writing courses at GWC. Further, they provide indirect evidence for the multiple measures placement model currently used at GWC in which the APS Writing test plays the primary role. Finally, these data support the argument that instructor ratings of student preparedness can be effective criteria in placement validation studies.

References

Anastasi, A. (1982). *Psychological Testing*. (5th Edition). New York: MacMillan Publishers.

Cabrera, A., Nora, A. & Castaneda, M. B. (1993). College Persistence: Structural Equations Modeling Test of an Integrated Model of Student Retention. *Journal of Higher Education*. 64(2), 123-139.

California Community Colleges. (1992). *Standards, Policies, and Procedures for the Evaluation of Assessment Instruments Used in the California Community Colleges*.

CAPP Associates. (1993). *Validation of College Board Tests: CPT, DTLS, DTMS, APS*. Report on statewide validity study submitted to the California Community College Chancellor's Office on behalf of the College Board.

Isonio, S. (1991). Implementation and initial validation of the APS English test. Golden West College, Huntington Beach, CA. (ERIC Document Number 345 781; Abstracted in *Resources in Education*, October, 1992).

Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*. 112(3), 527-535.

Matriculation Local Research Options Committee (1991). *Assessment Validation Project Local Research Options*. Sponsored by the California Community Colleges' Chancellor's Office.

Rasor, R. & Barr, J. (1993). *Refinement in assessment validation: Technicalities of dealing with low correlation and instructor grading variation*. American River College, October, 1993.

Woehr, D. J. & Cavell, T. A. (1993). Self-report measures of ability, effort, and non-academic activity as predictors of Introductory Psychology test scores. *Teaching of Psychology*, 20(3), 156-160.

APPENDICES

A: Descriptions of English writing courses targeted by the APS writing test at GWC

B: Directions for Instructor Ratings of Student Preparedness (with the rating scale to be used)

C: Table 1: Relationship between APS Writing test scores and instructor ratings of student preparedness

Appendix A

Descriptions of English writing courses targeted by the APS writing test at GWC

ENGLISH 9: BEGINNING ENGLISH (Precollegiate level)

This course is designed for native speakers of English who need basic work in sentence writing and paragraph development. One of a group of pre-English 100 courses. Course work includes functional grammar review, writing sentences, choosing a topic, narrowing it, and drafting a paragraph.

ENGLISH 10: WRITING ESSENTIALS (College level; non-transferable)

This course is designed to qualify the initially unprepared student for entry into English 100, and includes paragraph writing and grammar review.

ENGLISH 100: FRESHMAN COMPOSITION (College level; transferable)

Practice in the following: composition and revision of essays, critical thinking, critical reading and documentation. Recommended for liberal arts majors and those planning to transfer to a four-year college or university.

Appendix B

Directions for Instructor Ratings of Student Preparedness

Evaluate the preparedness of each student on your roster. This evaluation should be based on observation of students' demonstrated skills/abilities during the first weeks of the term. This would include students' ability to comprehend the material covered in the course which could be manifested in homework assignments, quizzes, and/or exams. Please do not take into account students' attendance, motivation, effort, or whether they submit their homework; the rating should be based strictly on academic preparation for the course.

Please use the following rating scale to evaluate levels of preparedness. Enter the number next to the student name on the roster:

Rating	Interpretation
1	Unprepared for the course. Probably should be enrolled in a lower course.
2	Marginally prepared for the course.
3	Adequately prepared for the course.
4	Well prepared for the course.
5	Exceptionally well prepared for the course. Possibly could be enrolled in a higher course.

Appendix C

Table 1

Relationship Between APS Writing Test Scores and Instructor Preparedness Ratings

Course	n	Sx	Sx-norm	r-obtained	r-corrected
English 9	112	3.17	6.23	.25	.45
English 10	152	3.38	6.23	.32	.53
English 100	48	3.19	6.23	.16	.30

Note:
n = Sample size
Sx = Sample standard deviation
Sx-norm = Normative standard deviation (from overall GWC dataset)
r-obtained = Correlation as calculated from data
r-corrected = Correlation corrected for range restriction

Acknowledgments

Robert Elmore of CAPP (Computerized Assessment & Placement Programs) performed the data matches and contributed substantively to the development of the rating scale. Patrick Kennedy, GWC English Department Chairperson, helped enlist the cooperation of the instructors.